DOCUMENT RESUME

ED 060 051 TM 001 136

| | |
|---|---|
| AUTHOR | Menaker, Shirley L.; And Others |
| TITLE | Prediction of Mental Health Through Computer Scoring of a Sentence Completion Test. |
| INSTITUTION | Texas Univ., Austin. Research and Development Center for Teacher Education. |
| PUB DATE | 66 |
| NOTE | 7p.; Paper presented at the American Psychological Association Convention, September 1966 |
| AVAILABLE FROM | R&D Center for Teacher Education, Ed. Annex, U. of Texas, Austin, Texas 78712 |
| EDRS PRICE | MF-$0.65 HC Not Available from EDRS. |
| DESCRIPTORS | Clinical Experience; *College Students; Computers; Correlation; Factor Analysis; Females; Grade Point Average; *Mental Health; Multiple Regression Analysis; Personal Adjustment; *Personality Assessment; *Predictive Measurement; Scoring; Sentence Structure; Tests; *Test Scoring Machines |
| IDENTIFIERS | Form 4 A; *Peck Veldman Sentence Completion Test |

ABSTRACT

An investigation was undertaken to determine how successfully an overall adjustment rating made by trained clinical raters could be predicted by means of a computer program for scoring adjustment which used only the data from a one word sentence completion instrument. Subjects for the study were 69 female college seniors at the University of Texas. The adjustment ratings on these subjects made by the clinical raters were based on an extensive assessment battery, which included the sentence completion instrument. (Author/CK)

ED 060051

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY

S. Menaker

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

# Prediction of Mental Health Through Computer Scoring of a Sentence Completion Test[1]

Shirley L. Menaker, Donald J. Veldman, and Robert F. Peck

The University of Texas

Sentence completion data from a one-word sentence completion form (the Peck-Veldman sentence completion, Form 4-A) were available from a sample of 69 female seniors enrolled in the college of education at the University of Texas. Two judges, who were experienced in clinical assessment, independently rated the subjects on 40 aspects of personality and behavior. The ratings were in the form of points on a seven-point rating scale. A rating of "overall mental health" was included. In making their ratings, the judges used the following data:

Biographical Information Form
School Situations Analyzer
2-D 90-item Sentence Completion (Multi-Word)
Peck-Veldman 90-item One Word Sentence Completion (OWSC)
Thematic Apperception Test (8 cards)

After completing their independent ratings, the raters met to discuss the cases and fill out a final rating scale form on each subject representing their agreement. These concensus ratings were then submitted to factor analysis of the 40 scales.[2] Uncorrelated factor scores were computed by means of regression equations for each factor.

It was found that most of the variation could be accounted for by two major factor dimensions: one labelled Rational Autonomy (22.78%) and

TM 001 136

[1]Paper presented at APA Convention, September, 1966 (Division 5).

[2]Principal components analysis, selecting factors with eigenvalues over 1.00 for orthogonal Varimax rotation.

the other Interpersonal Attitudes (18.95%), both of which factors loaded about equally on the overall mental health scale (loadings .61 and .58 respectively).

Loadings of the 40 scales on the first Principal Axis factor indicated that the central focus of the ratings was the overall mental health of the subjects. The rotated factor structure revealed that this central focus could really be considered the product of the two major factors mentioned above: Rational Autonomy and Interpersonal Attitudes. Factor scores based on this principal axis factor were thereafter used to provide a measure of overall mental health which would be more stable than simply the scale rating of mental health alone on each subject.

An investigation was then undertaken to determine how successfully this relatively stable overall mental health rating (as provided by the Principal Axis factor score) could be predicted by means of a computer program for scoring mental health which used only the data from the one-word sentence completion (OWSC). This scoring entailed the compilation of a mental health dictionary to be used by the computer.

To set up the dictionary, a response bank of compiled OWSC raw data for 1000 female sophomores enrolled in the college of education was used. This bank listed all responses to each of the 90 stems of the form separately and, within each stem, gave the frequency of each separate response. In compiling the "dictionary", the only responses included were those occurring with a frequency of greater than one, within the relevant stem, as obtained from this listing of the compiled raw data of the 1000 female sophomore students. This was done to maximize the predictive efficiency of the system with new samples, as idiosyncratic responses would have a low probability of being repeated in a different sample.

A clinical psychologist (who had not been involved in the previous ratings of the subjects) examined all responses to each stem in this listing and any responses judged to be clearly related to overall mental health (either positively or negatively) were pulled out for use in the scoring system. Three categories were set up: Positive (scored +1); Negative (scored -1); and Extreme Negative (scored -2).

When all stems had been examined, the words entering into this mental health dictionary were used to score the protocols of the 69 seniors. This scoring was done by the computer. As the data of the 69 protocols had previously been punched, once the dictionary was punched, raw scores were obtained for each subject reflecting the algebraic sum of the weights of words appearing in the individual protocol which also appeared in the mental health dictionary. In other words, every one of the 90 responses on an individual protocol was examined by the computer and, if the response also appeared in the mental health dictionary for that stem, its weight was added into a score for the subject.

These raw scores (giving a machine estimate of mental health from the OWSC) were correlated with the overall mental health rating based on the principal axis factor score outlined previously. The correlation was .5508. This correlation, while highly significant for the sample size, accounts for only about 30% of the variance. This means that the reduction of error in predicting the overall mental health rating with knowledge of the machine OWSC scores as against without knowledge of these scores is on the order of 16.5% ($E = 100\ (1-\sqrt{1-r^2})$, leaving much to be desired in the way of prediction.

At this point, it was noted that, while the overall mental health rating, the so-called criterion mental health score, correlated highly

with GPA on the subjects (r = .5535), a finding not surprising in itself, the machine rating was relatively independent of this variable. The correlation between GPA and the computer mental health score was .1282. This may be explained by the fact that the machine rating is based on weights assigned to individual words on the basis of their apparent relationship to deviant attitudes. The overall mental health rating, on the other hand, is based on the psychological opinions of human raters who are rating a much larger battery of projective data. These raters are correctly influenced by features such as grammar, spelling, word usage, organization, coherence, etc. when making their ratings of mental health.

Looking at the intercorrelation matrix of the three variables as presented in Table 1 (GPA, the criterion Mental Health Factor Score, OWSC Mental Health Machine Score), it was apparent that a substantial increase in the ability to predict the criterion mental health score would result from a combination of the variable dealing with the intellectual component (GPA) and that dealing with the measure of deviant attitude as expressed on the sentence completion and picked up by the computer scoring program.

Table 1

| Variable | $X^1$ | $X^2$ | $X^3$ |
|---|---|---|---|
| $X^1$ | | .5508 | .5535 |
| $X^2$ | .5508 | | .1282 |
| $X^3$ | .5535 | .1282 | |
| $M_x$ | 1.6694 | -.1304 | 1.7586 |
| $x$ | 17.9259 | 5.0128 | .3965 |

$X^1$ = Principal axis, overall mental health rating

$X^2$ = Mental health score from machine scoring of OWSC

$X^3$ = GPA

The "r" resulting from this multiple correlation was .7348, which after shrinkage remained .7250, a correlation coefficient approximating that which could be expected between skilled raters on a variable such as mental health.

Thus a prediction of overall mental health from two relatively independent scores which are easy to obtain (GPA and a computer-rated sentence completion mental health score) compared favorably with a rating of mental health achieved through the analysis of a large projective battery by two highly trained clinical raters. The results suggest that the administration of the Peck-Veldman One Word Sentence Completion and the construction of a mental health dictionary suitable for the population involved would offer an inexpensive means of identifying well adjusted and poorly adjusted people in the population with considerable accuracy. This should prove useful in large-scale selection of college subjects for a research design in which it was advisable to control for mental health.

To better evaluate the size of the correlation coefficient obtained in this study, it is necessary to keep in mind that it compares a computer scoring of mental health from sentence completion data with a measure of mental health obtained from clinical rater scoring of a battery of projective data. Thus, some of the difference in rating may result not merely from the limitations of the computer program but rather from the fact that a mental health rating from a sentence completion form alone may differ considerably from a rating obtained from a larger battery of projective data.

An interesting and related question is how well a computer scoring program for this sentence completion form can duplicate, not only a rating of mental health, but also the kinds of assessment statements on a number

of psychological variables that a psychologist would make from a reading of the sentence completion form. This would first involve a clinical rating of a number of personality constructs from sentence completion protocols, and then an investigation into the ability of a computer to duplicate these ratings, by means of separate dictionaries for each construct.

Such a program has been carried out and is being reported in another paper at these meetings.[3]

---

[3]Peck, R. F., Menaker, S. L., and Veldman, D. J. Automated Personality Assessment with Sentence Completion Data. Presented at APA Convention, September, 1966. (Division 5)